#2

10/048077

The Patent Office

INVESTOR IN PEOPLE

GB00/03006

EJU

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

REC'D 31 AUG 2000
WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 16 AUG 2000

Patents Form 1/77

Patents Act 1977
(Rule 16)

THE PATENT OFFICE
06 AUG 1999
RECEIVED

The Patent Office

1/77

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

06 AUG 1999

1. Your reference	HL 73059	
2. Patent application number *(The Patent Office will fill in this part)*	**9918657.9**	09AUG99 E468052-2 D02847 F01/7700 0.00 - 9918657.9
3. Full name, address and postcode of the or of each applicant *(underline all surnames)*	**FLYING NULL LIMITED** Harston Mill Harston Cambridge CB2 5NH	
Patents ADP number *(if you know it)* If the applicant is a corporate body, give the country/state of its incorporation	United Kingdom	07070519001.
4. Title of the invention **CODED LABEL INFORMATION EXTRACTION METHOD**		
5. Full name of your agent *(if you have one)*	Haseltine Lake & Co.	
"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*	Imperial House 15–19 Kingsway London WC2B 6UD	
Patents ADP number *(if you know it)*	34001	

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number	Country	Priority application number *(if you know it)*	Date of filing *(day/month/year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application	Date of filing *(day/month/year)*

8. Is a statement of inventorship and of right to a grant of patent required in support of this request? *(Answer "Yes" if:* *a) any applicant named in part 3 is not an inventor, or* *b) there is an inventor who is not named as an applicant, or* *c) any named applicant is a corporate body.* *See note (d))*	Yes

1 CODED LABEL INFORMATION EXTRACTION METHOD

2 INTRODUCTION

In certain types of human and machine-readable information-bearing label, a set of elements is used to represent the information contained in the label. This representation may be by varying the characteristics of the elements comprising the label, and also by the position in which the label elements are placed. Reading apparatus senses the characteristics and placement of the elements in order to decode the information contained within the label. The elements in the label are sequentially scanned, to discover the presence of the constituent elements and to measure their characteristics and position and thence to decode the information contained by the label.

This application describes label decoding methods by which information represented on such a label can be decoded. Typical label embodiments are where the label is manufactured from a plurality of magnetically active elements supported on a substrate. Information is coded by controlling the relative positions. Independent control of the physical properties of the elements, such as shape may also be used.

In alternative embodiments, information may be coded by controlling the relative angles of orientation of the easy axis of the elements, in some cases together with independent control of element shape, or other physical property.

3 DESCRIPTION OF THE INVENTION

3.1 Scanning Method

In certain types of label decoding system a label consisting of a set of elements is scanned, that is the elements are activated in turn by applying a signal to which the scanned element responds, and to which the reading apparatus is sensitive. Either or both of the means of generating the scanning signal, and the means of generating the received signal in the reading apparatus select a single elements in turn so that the signal from a particular element may be resolved from that of the other elements.

The means of scanning may be automatic by electronically scanning the interrogating signal and the receive sensitivity along the label or by an automatic mechanism causing the label elements to be interrogated in turn. Alternatively the scanning may be manually operated where the label is moved past the reading apparatus by hand, or alternatively the reading apparatus or part thereof may be moved past the label by hand.

3.2 Signal received from reader

Fig 1a shows a plan view of a label where the elements are made out of high permeability, low coercivity magnetic thin film material such as Atalante manufactured by IST(Belgium). In this case the elements are 1mmx1.5mm. The easy axis of the material is arranged to be parallel with the longitudinal axis of the label.

The reading antenna is passed by the label subjecting each element in turn to a region of null magnetic field formed by a permanent magnet arrangement. A superimposed high frequency field, in this case 6.348 kHz causes the material element to switch when it is present in the magnetic null. A receive coil is used to detect the harmonic signals resulting from the element switching, and the received harmonic signal is mixed to baseband and filtered prior to digitisation. Typical signals received from such a label, detecting the second harmonic of the element switching signals, when the read head was passed in proximity to the tag are shown in Fig 1b. The signal from a single element is shown in Fig 1c. This shows two half peaks of opposite sign with a zero crossing. It can be seen that the label signal approximates a superposition of single element responses at each position in which there is an element.

The signal is sampled with a fixed sample rate, and converted to a digital representation in an analogue to digital converter. The sample rate must be fast enough to capture the signal information in accordance with the Nyquist criterion, preferably with anti-aliasing filtering prior to the sampling operation.

The sample number of the centre of each element response depends upon the time at which the element passed the read head.

3.3 Velocity estimates from signal

In addition, the width of the element response depends upon the speed at which the element passed the read head. Information about the dynamics of the relative motion of the tag and the read head can be inferred from measuring the width of the signal. This can be achieved by measuring the width of the positive going and negative going peak associated with the element. It can also be achieved by measuring the separation between the positive and negative going peaks. This speed estimate is essential to estimate the dynamics of motion and to decode the information in the label reliably without compromising the available code space by adding regular features in the label.

In alternative implementations of the reading apparatus, the scanning motion of the magnetic null relative to the label may be achieved by electronically scanning using a set of excitation coils that are driven in a phased manner so as to cause the magnetic null to propagate in space. The null may be scanned at sufficient speed that the switching of the label elements may be detected directly from the receive coils without needing to superimpose any separate high frequency excitation. In this case the material element response coupled into the receiver is a single peak. Fig 1d shows 3 adjacent peaks of the fundamental signal from 3 magnetic elements recorded from a reading system employing electronic scanning. In this case the width

of the response from each magnetic element can be used to provide an estimate of the velocity of the scanning null as it passed over the element being scanned. This information can be used to estimate the detailed dynamics and errors of the system and also from misalignment of the label from the nominal read axis of the read head.

3.4 Amplitude Coded Signals

Information can be encoded into the label by varying the height of the material elements, that is the size of the element normal to the longitudinal axis of the label, or the width, that is the size of the element along the longitudinal axis.

By maintaining the element width constant, variations in element height predominantly cause a change in amplitude of the received signal as illustrated in Fig 2a for the mechanically scanned second harmonic system, and in Fig 2b for an electronically scanned system detecting the fundamental signal received from the label.

A set of linear test labels with 3 elements was manufactured with the element positions at 0mm, 1.5mm, 3mm along the longitudinal axis of the label. The element widths were 1mm and the height of the centre element varied between 1mm and 2mm. For each label, in Fig 2a, the peak position, the half peak width, the half peak separation and the amplitude difference between the half peaks were measured. The relative variation of these parameters of the central test element compared with the outer elements, which were maintained at constant height, were plotted. The measured central element amplitude (vampr), relative to the outer elements, shows a clear, near linear relationship to central element height. The measured central element relative position (vposr), width (vwidr) and separation (vsepr) parameters do not vary significantly with central element height, over and above the effects of experimental errors and imperfections in the manufacture of the test labels.

Similarly Fig 2b shows the relative variation of measured central element amplitude, position and width with element height when the test labels were measured in an electronically scanning reading system and shows a similar near linear relationship to the mechanically scanned reading method receiving second harmonic signals.

Element height can be used to code additional information with a detector based on amplitude in the receive processing. In the simplest case, two heights are used to modulate the elements in the label. This is illustrated in Fig 3a, which shows a label with elements of one of two heights. The corresponding received signal is shown in Fig 3b. The difference in received amplitude from elements of different heights can be observed. A typical design goal is to maximise the range at which a label may be read. A compromise is needed to maximise reading range for a 2-height label, at the same time as maximising the ability to discriminate the heights of individual elements. In order to discriminate the height of label elements reliably, certain of these elements may be constrained to be at a known height to act as a reference. In the example illustrated in 3a, the 2 end elements of the label are maintained at maximum height to act as a reference. A design rule may also constrain a certain

number of the inner elements to also be at maximum height. The reference elements may be at any a priori fixed position, or at a position determined by an a priori rule.

The half peak width, or half peak separation can be estimated and used to form an estimate of the velocity of the scanning, at the time when the element was scanned, in the case of a reader detecting the second harmonic of the label signal. Alternatively, in the case of a reader detecting the fundamental of the label signal, an estimate of the peak width can be used to form an estimate of the velocity of the scanning, at the time when the element was scanned. These estimates are typically K/w_est where K is some constant, and w_est is the estimate of width or half peak separation, as appropriate. The value of K typically depends on parameters of the reader, the element shape and the position of the element relative to the read antenna. Since K varies with element properties, to facilitate estimation of the scanning dynamics, the process to estimate K may be restricted to elements of the same type based on the amplitude of the element signals.

3.5 Peak finding algorithm

The first step in processing a received signal such as shown in Fig 1b or Fig 3b is to estimate the parameters of the peaks contained within the signal. This is achieved with the following steps:

1 The signal is delimited, that is the continuous received signal is broken up into a contiguous set of samples containing the received signal from one label, or a set of peaks that is likely to contain enough information to decode the label, in the case of non-registered repeating codes.

2 The signals from this delimited period are sorted into order of ascending magnitude. An indexed sort function, for example using the Quicksort method, such as indexx() (Numerical Recipies in C, second edition, ISBN 0 521 43108 5) is appropriate so that the original location of the point of any rank can be quickly found.

3 The positive half peaks are found by working down from the largest signal in magnitude in the sorted data, testing the signal points as follows. If the test signal point is a peak, that is its adjacent signals are lower in magnitude it is estimated as a positive half peak, and its amplitude is interpolated by interpolation of the peak and its neighbouring points.

4 From each positive half peak found, the signal is scanned forward and backwards to find the points at which it falls below a defined magnitude relative to the peak, this is interpolated to a fraction of a sample period A typical magnitude value is 0.8 of the peak value found. The half peak width is estimated as the difference in position between these points. The half peak position is estimated as the average of these two points. The half peak is rejected if it overlaps with any other pre-existing peak in the list, that is its width about its position interferes the width about any other peak position in the list. If it is not rejected, the half peak is appended to a list of half peaks.

5 The peak finding process stops when the magnitude of the next highest point falls below an absolute threshold, or when it falls below a defined fraction of the

largest magnitude signal, for example 0.3 of the largest magnitude signal, or if a limit for the number of peaks is reached.

6 The negative half peaks are found by working up in magnitude from the lowest signal in magnitude in the sorted data, testing signal points of increasing magnitude. If the test signal point is a negative peak, that is its adjacent signals are higher in magnitude it is estimated as a negative half peak, and its amplitude is interpolated by interpolation of the peak and its neighbouring points.

7 For each negative half peak, the signal is scanned forward and backwards to find the points at which it increases above a defined magnitude relative to the negative peak, this is interpolated to a fraction of a sample period. The half peak width is estimated as the difference in position between these points. The half peak position is estimated as the average of these two points. The half peak is rejected if it overlaps with any other pre-existing peak in the list, that is its width about its position interferes the width about any other peak position in the list. If it is not rejected by this test it is then appended to a list of half peaks.

8 The negative peak finding process stops when the magnitude of the next highest point falls above an absolute threshold, or when it falls above a defined fraction of the largest magnitude signal, or if a limit for the number of peaks is reached.

9 The resultant list of positive and negative half peaks is then sorted into range order. This process will interleave the consecutive half peaks.

10 Merging range consecutive pairs of half peaks in the range sorted half peak list forms a list of double peaks. The amplitude of the merged double peak is the difference in amplitude of two half peaks being merged. The position of the merged double peak is the average of the two half peak positions. The width of the merged double peak is the average of two half-peak widths. The separation of the merged double peak is the difference between the two half peak positions.

The results of this half peak finding process are illustrated in Fig 4a which shows the half peaks found when the signal shown in Fig 3b was processed, and shows the output of step 9 above. These peaks are shown with a cross at the located position and estimated amplitude. Fig 4b shows the magnitude of the double peaks in amplitude units. In this case the amplitudes are negative since the first half peak is a negative going peak. The difference in peak amplitude between the full height and reduced height elements in the label can be clearly seen. Fig 4c shows the double peak separations versus double peak position showing the variation across the scanning of the label.

In the case of a fundamental signal where there is just one peak per element, such as that shown in Fig 1d, a simpler process is may be used, missing steps 6 to 10 above since there is no need to find and merge double peaks.

10.1 Extraction of velocity function

There are N label elements positions whose centres are located at $X_1 .. X_N$ along the longitudinal axis of the label, and whose position we wish to estimate. It is assumed

that X_1 is at position 0.0 and that the final element, X_N is at the known position of corresponding to the length of the label L.

The peak list as described above contains estimates of the sample number, and hence times T_i at which the reader detected a double peak. For each element i, the time T_i. is simply the position in samples where the peak was measured multiplied by the sample period, which is the reciprocal of the sampling rate.

By using the estimated double peak width or double peak separation for each of the N peaks denoted W_i a velocity estimate for the i^{th} peak is $VE_i=1/W_i$ at time T_i.

The order M velocity function of time is modelled by a polynomial function. M is chosen to allow sufficient degrees of freedom to model the dynamics of label motion sufficiently accurately. It can range from 1 (constant acceleration model) upwards .

$$VF(t,M)=a_0+a_1t+a_2t^2+a_3t^3+....a_Mt^M$$

where {a} are a set of polynomial coefficients to be determined.

The squared error for velocity estimate VE_i at time T_i is $E_i^2=(VE_i-VF(T_i,M))^2$

The total sum of the squared errors is $ES(M)=E_1^2+ E_2^2+ .....+ E_N^2$ which is minimised by varying the M+1 parameters $a_0,a_1.....a_M$ This is achieved by the method of least squares to find the optimum parameters in the least squares sense.

Define the matrix X to contain the T_i values raised to a given power. For example column 0 row i of X contains T_i^0 and column M row 4of X contains T_4^M. Define the vector V to contain the velocity estimates VE_i. For example the element in row 4 of V contains VE_4

Some linear combination of the columns in X will be closest to the vector V in the least-squares sense. The coefficients of this linear combination will be the coefficients of the best-fitting polynomial. If A is the vector of length M+1 that holds the coefficients, the problem is to find the value for A that gives the best "near-solution" to the system of equations XA= V. If the columns of X are independent, the best choice for A is given by:

$$A=(X^TX)^{-1}(X^TV)$$

Where " ^{T} " represents matrix transpose and " $^{-1}$ " represents matrix inversion. The coefficient a_i is then the element from row i of A. This may be efficiently calculated by using methods such as LU decomposition, for examples the functions ludcmp() and lubksb() (Numerical Recipes in C, second edition, ISBN 0 521 43108 5)

This process is illustrate in Fig 4d where the velocity estimates W_i that were computed from the double peak separation data of Fig 4c are shown, those points marked "full data" contain the velocity estimates from all the peaks. In this example case, only the estimates from peaks with high amplitude, see Fig 4b, are selected for velocity fitting. These selected points are marked "Data points used in fit". If the label

known a priori to have elements of the same size then all elements would have been used. A least squares fit of order 2 was used and the continuous line shows the estimated velocity function as a function of sample number. A least squares fit of order 1, 2 or 3 is typically sufficient to model the velocity function with sufficient accuracy.

Error metrics are computed to allow the decoding algorithms to assess the quality of the read. A useful normalised velocity error metric, denoted evel, to quantify how well the velocity function fitted the data observed is defined. It is the sum of squares of the normalised error defined as :

$$evel=\mathrm{sum}(E_i^2/VE_i^2) \text{ for } i=1..N$$

The vector A could also be estimated by an iterative minimisation procedure such as Powell's method.

The above procedure can be repeated for varying model order M on the same data in order to find the best approximation. Preferably this is the solution with the smallest number of degrees of freedom, that fits the data read by the reader with the lowest error metrics.

10.2 Analytical integration and normalisation of tag length

The velocity fit function is $VF(t,M)=a_0+a_1t+a_2t^2+a_3t^3+....+a_Mt^M$ This may be analytically integrated to find the positional fit function

$$P(t,M) =Q+a_0t+(1/2)a_1t^2+(1/3)a_2t^3+....+(1/(M+1))a_Mt^{M+1}$$

where Q is an arbitrary constant of integration, set to zero.

We know that the position of label element 1 is at 0.0. We also know that the position of element N, the last element scanned is L, the length of the label. The final estimates of element position are:

$$XE_i = L\ .\ (P(X_i\ , M)-P(X_1,M))\ /\ (P(X_N\ ,M)-P(X_1,M))$$

10.3 Binning of data

There are N positional estimates XE_1 to XE_N from which N-1 gaps may be estimated.

Typically the gaps have the minimum gap subtracted, and are divided by the minimum gap, to give a real number representing the number of gap increments. The nearest integer is then taken to represent the gap for subsequent decoding. This can be expressed as:

The N-1 normalised gaps R_j are calculated by

$$R_j=(XE_{j+1}-XE_j-\text{Minimum\_Gap})/\text{Gap\_Increment} \qquad \text{for } j=1...N-1$$

The normalised gaps are then converted to their nearest integer representaton $G_j = \text{int}(R_j+0.5)$ for $j=1...N-1$, where int(x) means the truncated integer portion of x.

One useful error metric, denoted emax, is defined as the maximum magnitude of the difference between G_j and R_j. Another useful error metric, denoted eav, is the average magnitude of the difference between G_j and R_j.

Fig 4e shows the gap sequence data R_i computed from the data in Fig 4a-d, and shows the close correspondence or "snapping" to an integer grid, giving a high confidence that the gaps have been correctly measured. The gap sequence in this case is 3,2,2,1,7,0,0,2,1 .

10.4 Enumeration algorithms

Depending on the coding scheme used, the gap sequence is enumerated to a unique decoded value. Various codes exist that have different properties for example those codes allowing read direction to be determined, or those that allow non-registered operation, that is allowing the gap sequence to start at any position, for example a portion of a repeating gap sequence. The code may be determined a priori, or diagnosed by examining the read signals for the number of material elements present, or by estimating the sum of the gap numbers. It is preferable to determine this beforehand in order to minimise the chance of misinterpreting a given set of read data.

For certain codes, the number of elements in the label may be one of a range of possible values, with fixed label length in order to increase the number of possible codes.

The direction finding codes are preferably used where amplitude coding is used to allow the amplitude information to be associated with the correct element.

10.5 Amplitude threshold estimation of elements

Where it is desired to detect the size of the received element signal, for example when the elements may be of two or more sizes, reference elements may be placed in the set of elements used to make the label. In one embodiment, the end elements are constrained to be the largest element size to act as a reference. Other embodiments may place reference elements that have some unique characteristic, or may be identified by an a priori rule.

The amplitudes are measured and used to compute normalised amplitudes. This is because the reference elements may be significantly different in magnitude. This may be due to the scanning process varying the distance or angle of the read head from the label during the read, or other factors.

The estimates of element position XE_i may be used for linear interpolation of the reference amplitude along the label. The reference amplitude predicted at the estimated position of the element divides the measured element amplitudes to normalise the amplitudes. The normalised amplitudes should be nominally 1.0 for an element the same size as the reference element.

Threshold levels are set to provide discrimination between the possible element amplitudes and to classify the amplitude of the element into one of an alphabet of values. For example if there are two values then the amplitude is either high or low. For example a half height element may have an amplitude, on average, half the reference amplitude ie 0.5. In this case, the threshold level should be set to approximately 0.75 in order to determine if the element was full height or half height.

Preferably a code is used which enables the direction of reading of the label to be determined by the sequence of gaps that resulted. In this case the elements can be uniquely identified and it is known which element of the label to associate with the decoded amplitude information.

In the case of the label data shown in Fig 4b, denoting the amplitudes of the signals from the elements as "H" for high amplitude and "L" for low amplitude, the elements are categorised as H,L,H,L,H,L,H,L,L,H. A larger alphabet of elements could be used by using a label with more possible shapes, for example. The categorisations are then converted to decoded data.

A typical scheme allows the inner elements only to contain data, ignoring the end elements, which are references, and maps H to a logical 0 and L to a logical 1 so that the binary data word from the example data would be 10101011. In this scheme up to 256 amplitude codes can be represented from 8 inner elements. Preferably the codes used are constrained to ensure that a certain number of elements in the label, for example 4 are at full height to facilitate velocity estimation. If this constraint is used, the number of amplitude codes is reduced by 37 to 219 in this case.

An amplitude error metric eamp is defined which is the normalised amplitude difference between the largest half height element diagnosed to the smallest full height element diagnosed. If there is no half-height element diagnosed, eamp is defined as the smallest full height element amplitude. In a label containing different sized elements, a large value for this metric suggests a clear discrimination in amplitude between elements of different types. Smaller values indicate increasing risk that an element may be categorised as the wrong amplitude.

10.6 Error Metrics

The error metrics defined: evel, emax, eav and eamp can be used to determine the read quality. evel, emax and eav are preferably small and eamp high. Independent thresholds for these quantities are independently established. These thresholds are based on measured values for many labels in many read operations over the range of reading ranges and scanning dynamics expected in an application. This is done to

minimise the probability of a wrong code at the same time as maintaining a high probability of a correct read.

During the read decoding process, if the computed error metrics are higher than the threshold value in the case of evel, emax and eav then the read is rejected. . If the error metric is lower than the threshold in the case of eamp then the read is rejected.

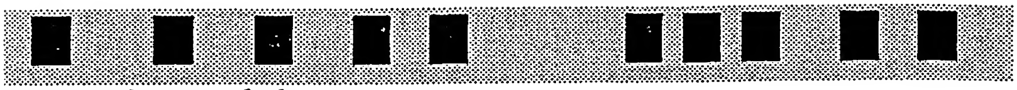

Fig 1a: Linear Label

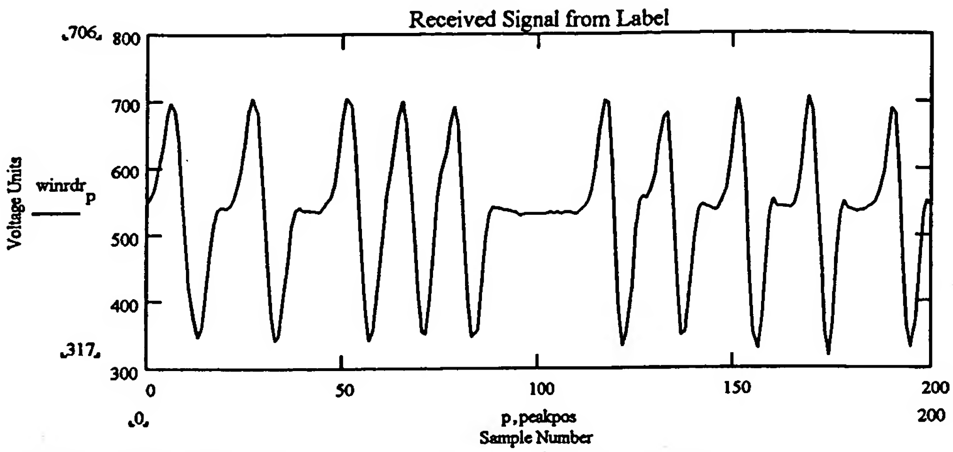

Fig 1b: Typical received second harmonic signal from label

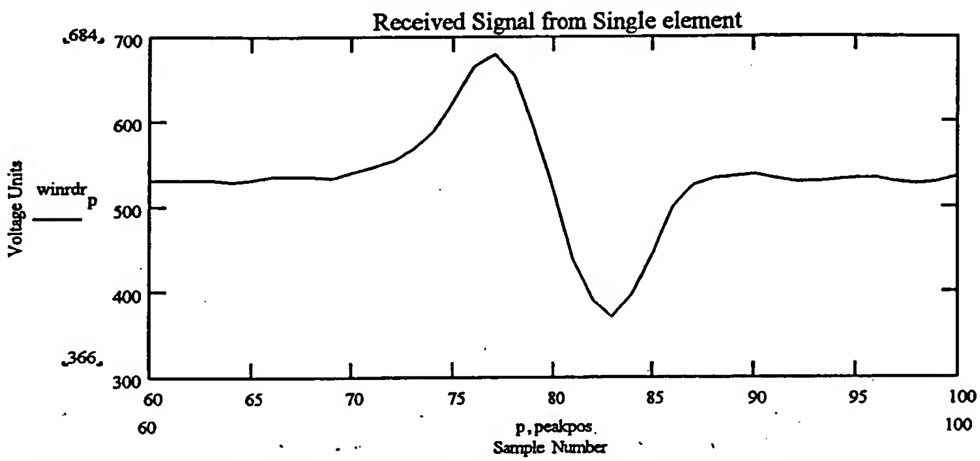

Fig 1c: Typical received second harmonic signal from element of label

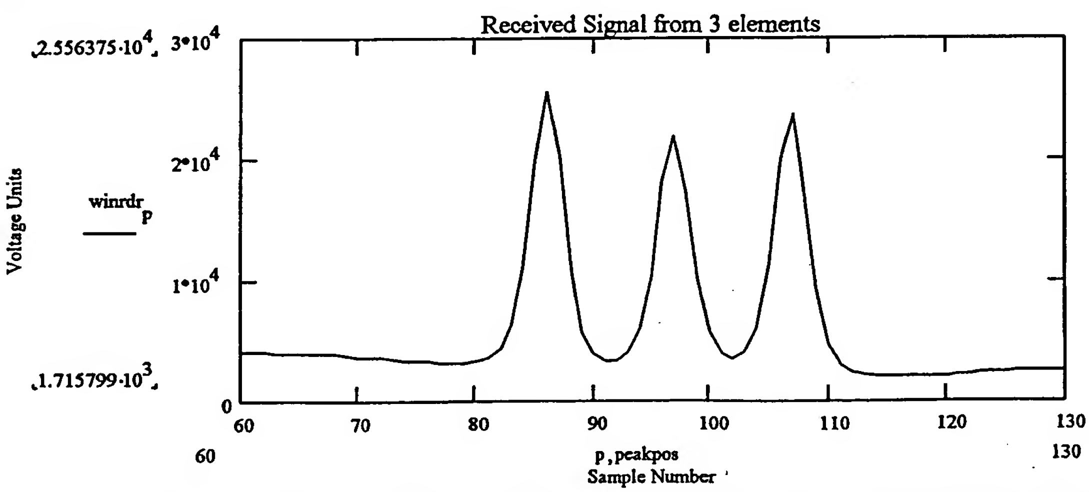

Fig 1d: Received fundamental signal from 3 elements in electronically scanned reader

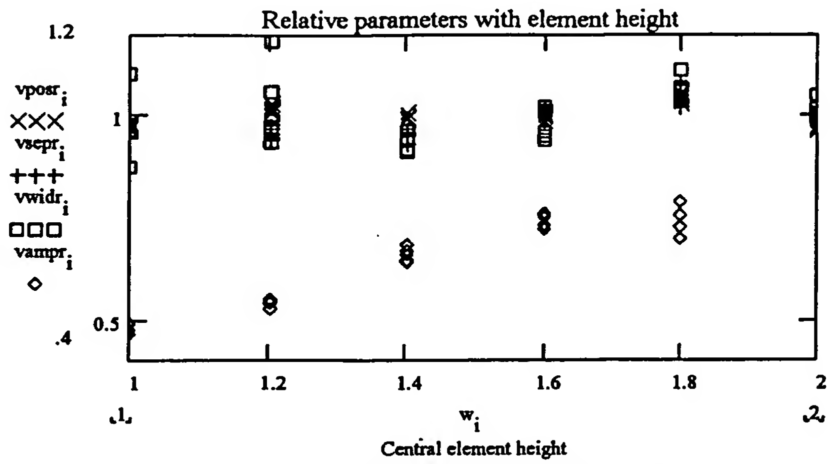

Fig 2a: Variation of amplitude with element height, second harmonic signal

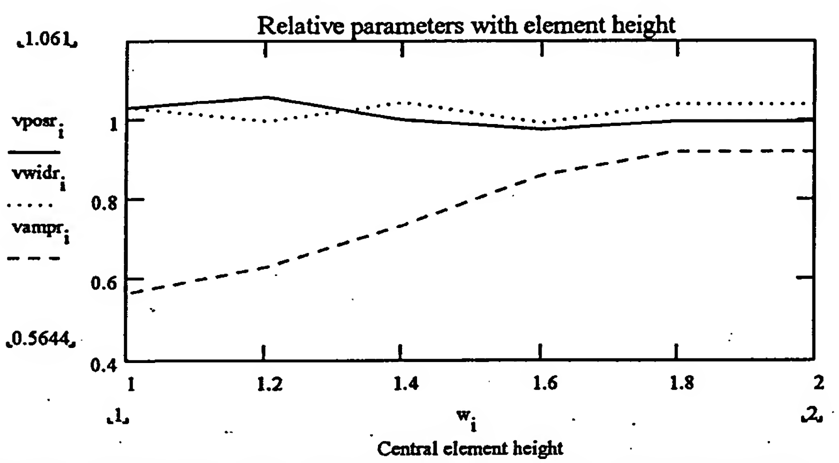

Fig 2b: Variation of received amplitude with element height – fundamental signal

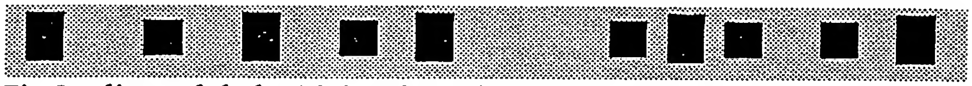

Fig 3a: linear label with height coding

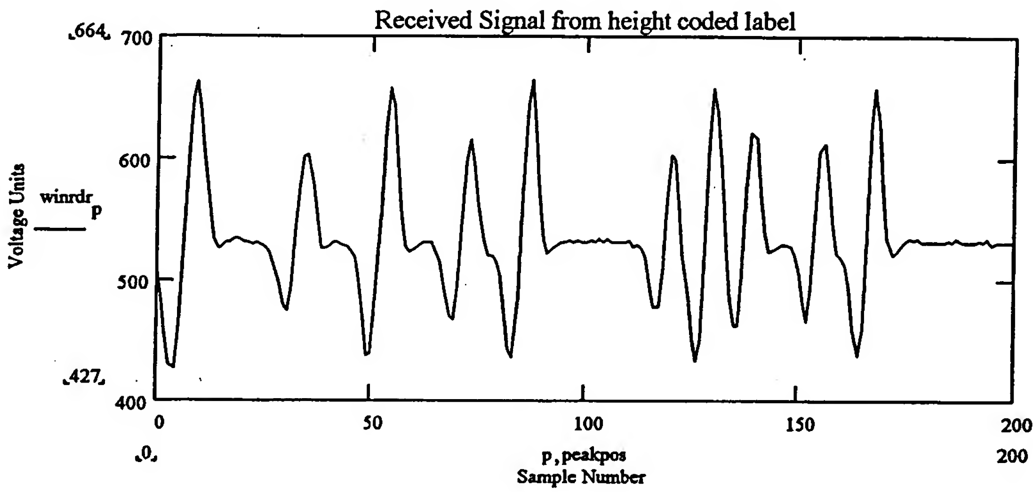

Fig 3b: Receive signal from linear label with height coding

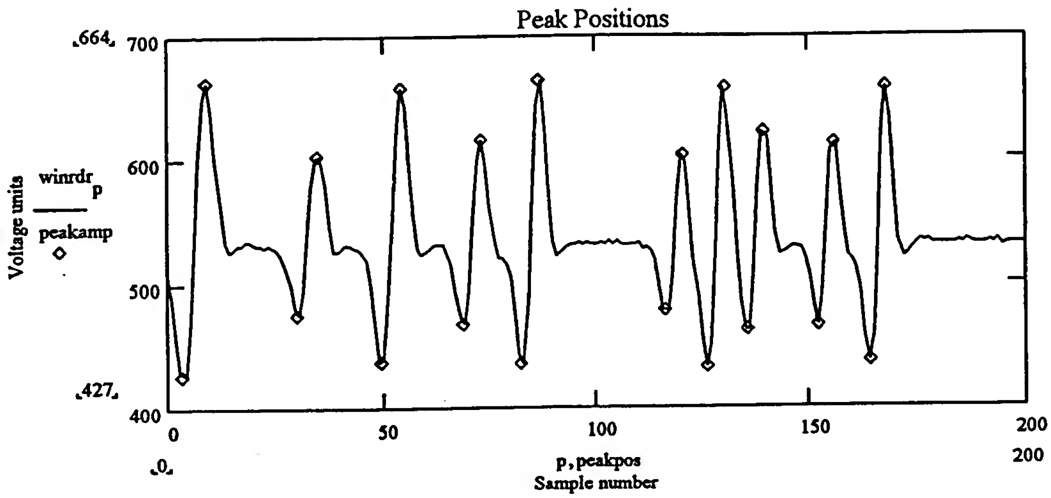

Fig 4a: Detected half peak positions

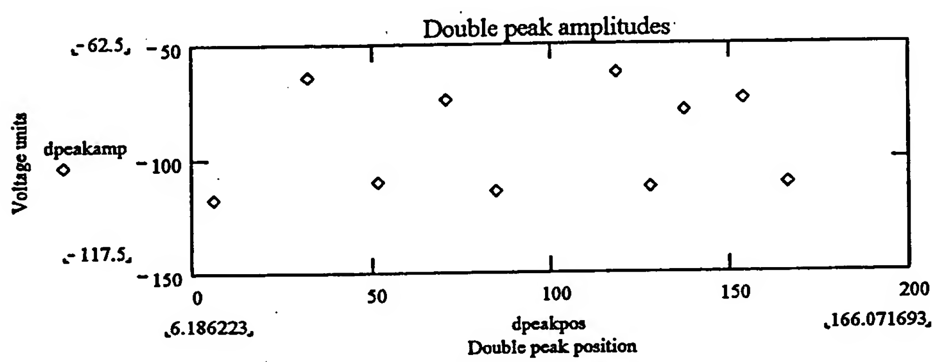

Fig 4b: Double peak amplitudes versus time
Element amplitudes categorised : H,L,H,L,H,L,H,L,L,H

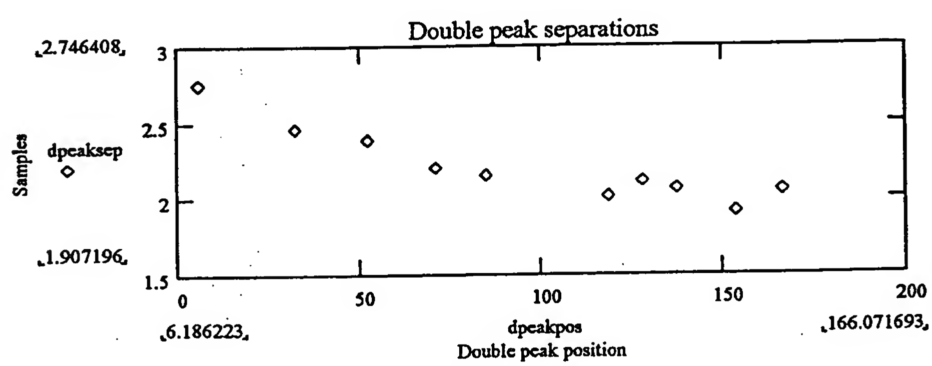

Fig 4c: Double peak separation versus time

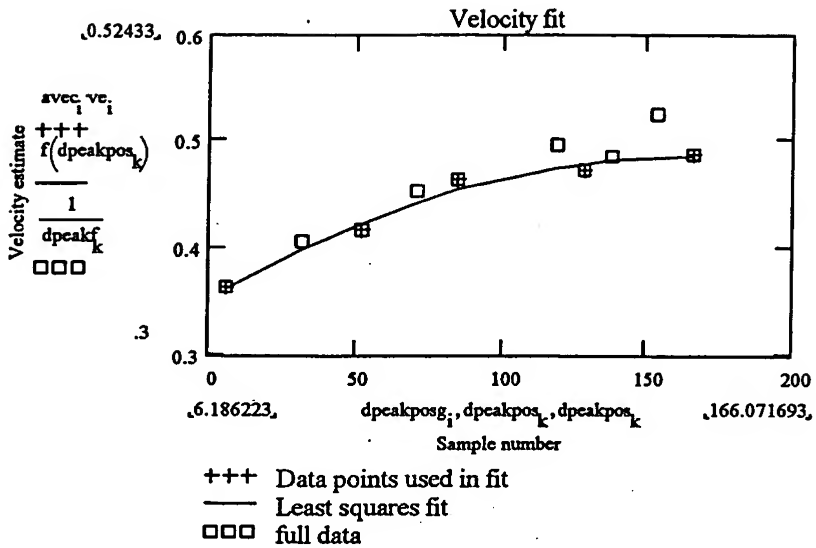

Fig 4d: Velocity fit to selected points versus time

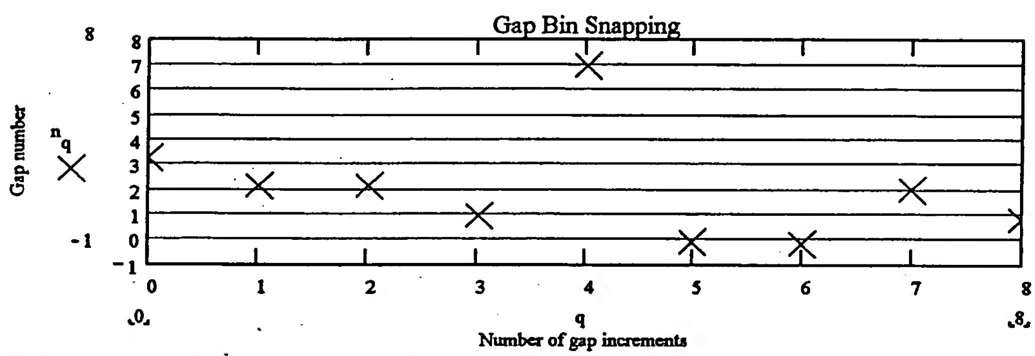

Fig 4e: Resultant fit of Gap number to integer grid
Nearest integers to gap number is 3,2,2,1,7,0,0,2,1